HISTORY 4816B / 9877B
Digital Research Methods
Fall/Winter 2021-22
4 hours per week
Asynchronous online course

Instructor: Professor William J. Turkel

Office Hours: Online by appointment
Department of History, Office: Lawson Hall 2267
Email: wturkel@uwo.ca

This is a draft syllabus. Please see your course OWL site for the final syllabus.

Course Description:

In this course students will learn how historical content is produced, presented and published online; how to find and evaluate digital primary and secondary sources; and how to use computational techniques to work with digital resources. No previous background in the subject area is required.

Prerequisite(s):

Registration in any academic program at the second-year level or above.

Unless you have either the prerequisites for this course or written special permission from your Dean to enroll in it, you may be removed from this course and it will be deleted from your record. The decision may not appealed. You will receive no adjustment to your fees in the event that you are dropped from a course for failing to have the necessary prerequisites.

Antirequisite(s):

Digital Humanities 2130A/B, Digital Humanities 3902F/G, the former History 2816F/G.

Course Syllabus:

Research of all kinds now crucially involves the acquisition and use of digital sources, both primary and secondary. In this course, you will learn to find, harvest, manage, excerpt, cluster and analyze digital materials throughout the research process, from initial exploratory forays through the production of an electronic article or monograph which is ready to submit for publication. The digital sources that we work with include texts, images, maps, code, social media, spreadsheets, databases and linked open data. The techniques that we will employ are currently used by relatively few humanists and social scientists—although they have become much more crucial in an age of pandemic lockdowns. Computation won't magically do your research for you, but it will make you much more efficient. You can focus on close reading, interpretation and writing, and use machines to help you find, summarize, organize and visualize sources.

Prior Experience

The only thing that you need to do well in this course is a willingness to learn new things and the perseverance to keep working when you're confused or when you realize that you could spend a lifetime learning about the topics and technologies that we will cover in class, and still not master them all. Students will come into the course with very different levels of experience and expertise. Some, probably most, will be familiar only with the rudiments of computer and internet use. A few may already be skilled programmers.

Workload

This course requires that you spend at least a bit of time each day (say 20-30 minutes) practicing your new skills. It's a lot like learning a new language, learning to play a musical instrument or going to the gym. At first it is going to be hard but be patient with yourself and ask a lot of questions. With daily practice, you will soon find ways to do your research and coursework faster and more efficiently. If you can't commit to regular practice, however, you should probably not take this course. The techniques that you learn in this class build cumulatively week-by-week, and there are a number of evenly spaced assignments of equal value.

Learning Outcomes

At the conclusion of this course, you will

- Be aware of a wide variety of different kinds of digital sources and strategies for making use of them computationally
- Be familiar with a range of research approaches in digital humanities, computational social science and related disciplines
- Be acquainted with methodological applications of artificial intelligence
- Be more familiar with using computer code as a medium of expression

Course Materials:

To get the most out of this class, you will need a Windows, Mac or Linux laptop or desktop. A tablet or netbook will probably not be sufficient. If you have any concerns, email me.

You should purchase a **student desktop license** for Wolfram Research's *Mathematica* software. (Don't let the name scare you, you won't need any particular training in mathematics to do well in the class).

http://www.wolfram.com/mathematica/pricing/students-individuals.php

You can purchase the license for the semester, for the year or permanently. If you are unwilling or unable to purchase the software, please do not take this course.

The required text for the course is freely available online:

William J. Turkel. *Digital Research Methods with Mathematica*, 2nd rev ed. (2020).
https://williamjturkel.net/digital-research-methods-with-mathematica/

Methods of Evaluation:

Assignment	*Worth*	*Handed Out*	*Due Date*	*Zero Date*
01	20%	Jan 11	Jan 18	Jan 25
02	20%	Jan 25	Feb 01	Feb 08
03	20%	Feb 08	Feb 15	Feb 28
04	20%	Feb 28	Mar 07	Mar 14
05	20%	Mar 14	Mar 21	Apr 01

Each assignment has a suggested due date and, a week later, a zero date. If you hand in your assignment before the zero date, you will receive full credit for your work. After the zero date, it will be worth nothing.

If you are unable to meet a course requirement due to illness or other serious circumstances, you must provide valid medical or other supporting documentation to Academic Counseling as soon as possible.

Accommodation for missed tests/midterms, including Self Reported Absences (SRA):

Students with an approved absence from an in-class test will be required to write a makeup test.

Students should be aware that the make-up test will not necessarily be in the same format, be of the same duration, or cover the same material as the original test.

Online courses will have online makeup tests.
In person classes have scheduled makeup tests scheduled at the following times:

Fall Term

- Tests scheduled before November 1 – the makeup will take place November 8 at 9:30am.
- Tests scheduled between November 2 and December 7 – the makeup will take place December 8 at 9:30pm.

Winter Term

- Tests scheduled before February 19 – the makeup will take place February 28 at 9:30am.
- Tests scheduled between February 20 and March 31 – the makeup will take place April 1 at 12:30pm

No other make-up opportunities will be provided. Students who fail to write a makeup test in the designated time slots will receive a grade of zero.

Course professors may not be available to respond to questions during the makeup test slots.

Students should be aware that when they have submitted an SRA for one test, they are not permitted to write a test or complete an assignment for another course during the period covered by the SRA. Failure to observe this regulation will result in the cancellation of the SRA and the possible application of late penalties.

Accommodation for missed assignment deadlines with a Self Reported Absence:

If a student reports a SRA for an assignment (i.e. an essay) the new due date will be 48 hours after the SRA was submitted.

Course Schedule and Readings:

- **Jan 04. Introduction to *Mathematica***. Interacting with notebooks
- **Jan 06. Reading Code**. Word frequency, word clouds and stopwords.
- **Jan 11. Computable Knowledge**. Entities, tables, timelines and maps.
- **Jan 13. Text Content**. Mathematica notebooks and expressions, strings and natural language processing.
- **Jan 18. Data Structures**. Lists, associations and datasets.
- **Jan 20. Reusing Code**. Defining and developing functions, keyword in context (KWIC).
- **Jan 25. Networks**. Metadata, matrices and social network analysis.
- **Jan 27. Indexing and Searching**. Pattern matching, topic classification and term distribution.
- **Feb 01. Geospatial Analysis**. Geographic information: raster, vector and attribute data.
- **Feb 03. Images**. Computer vision, face detection, feature extraction and image mining.
- **Feb 08. Page Images**. Optical character recognition (OCR), figure extraction and classification.
- **Feb 10. Crawling**. Browser automation, batch downloading, web archives and WARC files.
- **Feb 15. Linked Open Data**. Resource description framework (RDF), SPARQL queries and endpoints, JSON-LD.
- **Feb 17. Markup Languages**. Scraping and parsing, XML, really simple syndication (RSS) and text encoding initiative (TEI).
- **Feb 22 & Feb 24**. WINTER STUDY BREAK
- **Mar 01. Studying Societies**. Computational social science, search data, social media and social networks.
- **Mar 03. Extracting Keywords**. Information retrieval, term frequency-inverse document frequency (TF-IDF) and rapid automatic keyword extraction (RAKE).
- **Mar 08. Word and Document Vectors**. Feature extraction, dimension reduction, word embeddings and global vectors.
- **Mar 10. Citations**. References, web services, bibliographic linked open data and citation networks.
- **Mar 15. Natural Language**. Multilingual analysis, computational linguistics and sentiment analysis.
- **Mar 17. Web Services**. Entity networks, publication search, dashboards, manipulating JSON.
- **Mar 22. Databases**. Parts, selections and transformations, computations and querying, relations.
- **Mar 24. Measuring Images**. Photogrammetry, georectification, handwriting and facial 3D reconstruction.
- **Mar 29. Machine Learning**. Unsupervised clustering, classify, predict and transfer learning.

Additional Statements:

Please review the Department of History Course Must-Knows document, https://www.history.uwo.ca/undergraduate/Docs/Department%20of%20History%20Course%20Must-Knows.pdf, for additional information regarding:

- Academic Offences
- Accessibility Options
- Medical Issues
- Plagiarism
- Scholastic Offences
- Copyright
- Health and Wellness